

(19)



JAPANESE PATENT OFFICE

PATENT ABSTRACTS OF JAPAN

(11) Publication number: **2001100973 A**

(43) Date of publication of application: **13.04.01**

(51) Int. Cl.

G06F 3/16
G06F 17/24
G10L 15/00
G10L 15/22
G10L 15/28

(21) Application number: **11282091**

(71) Applicant: **NEC SOFTWARE KOBE LTD**

(22) Date of filing: **01.10.99**

(72) Inventor: **TOHARA HAJIME**

(54) **METHOD AND DEVICE FOR COMPUTER
DICTATION AND RECORDING MEDIUM
STORING COMPUTER DICTATION PROGRAM**

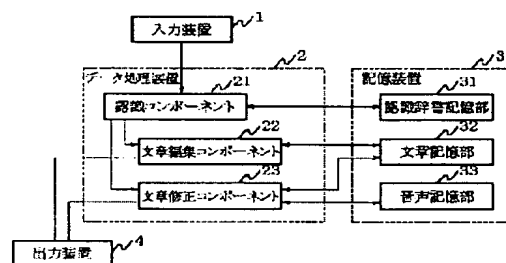
correction is received.

COPYRIGHT: (C)2001,JPO

(57) Abstract:

PROBLEM TO BE SOLVED: To improve the voice recognition rate by reducing the consumption of the storage area of a dictation device.

SOLUTION: A recognition component 21 receives only a navigation for editing or dictation inputted in voice from an input device in dictation mode and also receives only a navigation for correction inputted in voice from the input device in correction mode. A document editing component 22 performs a process corresponding to the kind of the navigation for editing when the voice input is the navigation for editing, and then stores the result in a document storage part 32 and also outputs it to an output device 4. A document correcting component 23 outputs the character string of the dictation to the document storage part 32, voice information to a voice storage part 33, and the character string to the output device when the voice input is the dictation, and performs a process corresponding to the kind of the navigation for correction when the navigation for



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2001-100973

(P2001-100973A)

(43) 公開日 平成13年4月13日 (2001.4.13)

(51) Int.Cl. ⁷	識別記号	F I	テ-マコード [*] (参考)
G 0 6 F 3/16	3 2 0	G 0 6 F 3/16	3 2 0 H 5 B 0 0 9
17/24		15/20	5 5 4 N 5 D 0 1 5
G 1 0 L 15/00		G 1 0 L 3/00	5 5 1 B 9 A 0 0 1
15/22			5 6 1 E
15/28			5 7 1 J

審査請求 有 請求項の数14 O L (全 11 頁)

(21) 出願番号 特願平11-282091

(22) 出願日 平成11年10月1日 (1999. 10. 1)

(71) 出願人 000192545

神戸日本電気ソフトウェア株式会社

兵庫県神戸市西区高塚台5丁目3番1号

(72) 発明者 戸原 一

兵庫県神戸市西区高塚台5丁目3番1号

神戸日本電気ソフトウェア株式会社内

(74) 代理人 100082935

弁理士 京本 直樹 (外2名)

Fターム(参考) 5B009 KB01 KB05 QB18

5D015 KK03 LL05 LL10

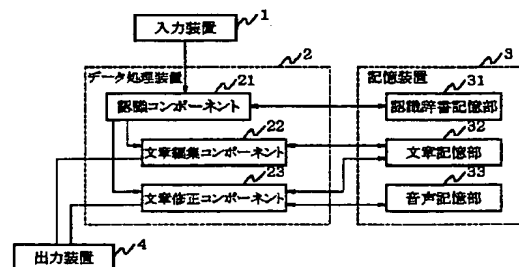
9A001 HH11 HH15 HH17 KK42

(54) 【発明の名称】 コンピュータディクテーション方法及びコンピュータディクテーション装置及びコンピュータディクテーションプログラムを格納した記録媒体

(57) 【要約】

【課題】 ディクテーション装置の記憶領域の使用量を低減させ、音声認識率の向上を図る。

【解決手段】 認識コンポーネント21は、ディクテーションモードにある時入力装置から音声入力された編集用ナビゲーションまたはディクテーションのみを受け取り修正モード時には入力装置から音声入力された修正用ナビゲーションのみを受け取る。文章編集コンポーネント22は、音声入力編集用ナビゲーションの場合は編集用ナビゲーション種別に応じた処理を行いその結果を文章記憶部32に記憶するとともに出力装置4に出力する。文章修正コンポーネント23は、音声入力ディクテーションの場合はその文字列を文章記憶部32に、音声情報を音声記憶部33に、文字列を出力装置に出力し、修正用ナビゲーションを受け取った場合は修正用ナビゲーション種別に応じた処理を行う。



【特許請求の範囲】

【請求項 1】 コンピュータディクテーション方法において、一発声分のディクテーション入力ごとにディクテーションモードから修正モードへの切り換えを自動的に行うことを特徴とするコンピュータディクテーション方法。

【請求項 2】 前記ディクテーションモードではディクテーションの音声入力と文字列の置換を行わない範囲での修正ができる編集用ナビゲーションの音声入力のみが可能であることを特徴とする請求項 1 記載のコンピュータディクテーション方法。

【請求項 3】 前記ディクテーションモードで音声により入力された文章は予め備えた認識辞書記憶部における辞書を用いて前記音声に対応した文字列に変換され以前に入力された文章と異なったフォーマットで出力装置に表示することにより新しく入力された文章が区別できるようにすることを特徴とする請求項 1 及び 2 記載のコンピュータディクテーション方法。

【請求項 4】 前記ディクテーションモードで前記編集用ナビゲーションとして音声入力された単語は予め備えた認識辞書記憶部における辞書を用いて音声認識され、認識された前記音声に対応した前記編集用ナビゲーションの処理が行われることを特徴とする請求項 2 記載のコンピュータディクテーション方法。

【請求項 5】 前記修正モードではディクテーション入力された前記一発声分の文章に対して単語の候補の変更・決定及び文章の確定などの操作を行う修正用ナビゲーションの音声入力のみが可能であることを特徴とする請求項 1 記載のコンピュータディクテーション方法。

【請求項 6】 前記修正用ナビゲーションとして音声入力された単語は予め備えた認識辞書記憶部における辞書を用いて音声認識され、認識された前記音声に対応した前記修正用ナビゲーションの処理が行われることを特徴とする請求項 5 記載のコンピュータディクテーション方法。

【請求項 7】 前記修正用ナビゲーションとして入力された音声単語候補の表示を指示する音声入力の場合は、前記ディクテーションモードで入力された音声であって予め用意された音声記憶部に記憶された音声情報を用いて単語候補の一覧を作成しこれを出力装置に出力することを特徴とする請求項 6 記載のコンピュータディクテーション方法。

【請求項 8】 前記修正用ナビゲーションとして入力された前記音声単語が文章修正終了を指示する音声入力の場合は、前記音声記憶部に記憶された音声情報を削除し、前記出力装置に表示された文章のフォーマットを以前に入力された文章のフォーマットと同一にすることを特徴とする請求項 3 及び 6 記載のコンピュータディクテーション方法。

【請求項 9】 コンピュータディクテーション装置にお

いて、ディクテーションモードにある時入力装置から音声入力された編集用ナビゲーションまたはディクテーションのみを受け取り修正モード時には前記入力装置から音声入力された修正用ナビゲーションのみを受け取る認識手段と、前記認識手段が前記編集用ナビゲーションを受け取った場合は改行、カレット移動などの前記編集用ナビゲーション種別に応じた処理を行いその結果を文章記憶部に記憶するとともに出力装置に出力する文章編集手段と、前記認識手段が前記ディクテーションを受け取った場合はその文字列を前記文章記憶部に記憶しその音声情報を音声記憶部に記憶し前記文字列を出力装置に出力し前記修正用ナビゲーションを受け取った場合は前記修正用ナビゲーション種別に応じた処理を行いその結果を前記出力装置に出力する文章修正手段とを含んで構成されることを特徴とするコンピュータディクテーション装置。

【請求項 10】 前記ディクテーションに対しては入力された音声に対応した文字列が、前記ナビゲーションに対しては対応する動作およびその動作が前記編集用ナビゲーションであるか前記修正用ナビゲーションであるかの情報が記憶されている認識辞書記憶部を予め具備し、前記認識手段は、与えられた前記音声入力の前記認識辞書記憶部に前記ナビゲーションとして登録されていればその情報により前記編集用ナビゲーションまたは前記修正用ナビゲーションであると判断し前記ナビゲーションとして登録されていなければ前記ディクテーションと判断することを特徴とする請求項 9 記載のコンピュータディクテーション装置。

【請求項 11】 前記文章修正手段は、受け取った前記ディクテーションの文章を前記出力装置に出力する時、以前に入力された文章とフォーマットを変えて表示することにより、新しく入力された文章が区別できるようにしさらにこの新しく入力された文章に対しての修正が可能な前記修正モードへの切り換えを行うことを特徴とする請求項 9 記載のコンピュータディクテーション装置。

【請求項 12】 前記文章修正手段は、前記修正モードで入力された前記修正用ナビゲーションが単語候補の表示を指示する音声入力の場合は、前記認識手段に前記音声記憶部に記憶された該当単語の前記音声情報を渡して候補一覧の要求を行いその結果を前記出力装置に出力することを特徴とする請求項 9 記載のコンピュータディクテーション装置。

【請求項 13】 前記文章修正手段は、前記修正モードで入力された前記修正用ナビゲーションが文章修正終了を指示する音声入力の場合は、前記音声記憶部に記憶された音声情報を削除し、前記出力装置に表示されたフォーマットを以前に入力された文章のフォーマットと同一にすることを特徴とする請求項 9 及び 11 記載のコンピュータディクテーション装置。

【請求項 14】 データ処理装置にあって請求項 9 から

請求項13における少なくとも1つの請求項における前記認識手段及び前記文章編集手段及び前記文章修正手段と同等の動作を実行するコンピュータディクテーションプログラムを格納した記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明はコンピュータディクテーションシステムに関し、特にナビゲーション機能を有するコンピュータディクテーション方法及びコンピュータディクテーション装置及びコンピュータディクテーションプログラムを格納した記録媒体に関する。

【0002】

【従来の技術】従来のディクテーション機能およびナビゲーション機能を有する音声認識システムを用いたコンピュータディクテーションシステムでは、入力された全ての文章に対して、改行、空白挿入、文章削除、カレット移動などの文章編集や、各単語の候補表示および修正をすることができた。このため、常にディクテーション、編集用ナビゲーション、修正用ナビゲーションが可能である必要があり、音声情報を保持しておく必要があった。

【0003】

【発明が解決しようとする課題】上述した従来のコンピュータディクテーションシステムでは、常にディクテーション、編集用ナビゲーション、修正用ナビゲーションを可能にするため、音声情報を保持しておく必要があった。そのため、長い文章を入力する際には、音声情報を保持しておくことが記憶領域を圧迫するという欠点があった。

【0004】また、ユーザが編集用または修正用のナビゲーション入力をしたつもり音声入力ディクテーション入力と誤認識されたり、あるいはその逆といったことが起こり得た。例えば、修正用ナビゲーションとして、単語の候補一覧を表示する「候補の表示」が登録されている音声認識システムがあるとする。今ユーザが、文章中の単語を変更するためにマイクに向かって「候補の表示」と発声したとし、さらに、「候補」の発音が不明瞭なために、システムが「広報の表示」と誤認識した場合を考える。「広報の表示」が編集用または修正用ナビゲーションとして登録されている場合には、システムはその動作を実行し、ユーザが望んだ以外の動作することになる。また、ナビゲーションとして登録されていない場合には、ディクテーションとして認識され、現在のカレット位置に「広報の表示」の文字列が挿入されることになる。この場合、ユーザは単語を修正するためにナビゲーションを行おうとしたにも関わらず、実際には余分な文字が挿入され、さらに修正が必要な結果となる。このように、従来のシステムでは、ディクテーション、編集用ナビゲーション、修正用ナビゲーションが常に可能なために、誤認識による操作性の低下を招くとい

う欠点もあった。

【0005】本発明は、以上のような問題点を解決する方法を提供する。

【0006】

【課題を解決するための手段】本願の第1の発明は、コンピュータディクテーション方法において、一発声分のディクテーション入力ごとにディクテーションモードから修正モードへの切り換えを自動的に行うことを特徴とする。

【0007】本願の第2の発明は、第1の発明の前記ディクテーションモードではディクテーションの音声入力と文字列の置換を行わない範囲での修正ができる編集用ナビゲーションの音声入力のみが可能であることを特徴とする。

【0008】本願の第3の発明は、第1及び第2の発明の前記ディクテーションモードで音声により入力された文章は予め備えた認識辞書記憶部における辞書を用いて前記音声に対応した文字列に変換され以前に入力された文章と異なったフォーマットで出力装置に表示することにより新しく入力された文章が区別できるようにすることを特徴とする。

【0009】本願の第4の発明は、第2の発明の前記ディクテーションモードで前記編集用ナビゲーションとして音声入力された単語は予め備えた認識辞書記憶部における辞書を用いて音声認識され、認識された前記音声に対応した前記編集用ナビゲーションの処理が行われることを特徴とする。

【0010】本願の第5の発明は、第1の発明の前記修正モードではディクテーション入力された前記一発声分の文章に対して単語の候補の変更・決定及び文章の確定などの操作を行う修正用ナビゲーションの音声入力のみが可能であることを特徴とする。

【0011】本願の第6の発明は、第5の発明で前記修正用ナビゲーションとして音声入力された単語は予め備えた認識辞書記憶部における辞書を用いて音声認識され、認識された前記音声に対応した前記修正用ナビゲーションの処理が行われることを特徴とする。

【0012】本願の第7の発明は、第6の発明の前記修正用ナビゲーションとして入力された音声単語候補の表示を指示する音声入力の場合は、前記ディクテーションモードで入力された音声であって予め用意された音声記憶部に記憶された音声情報を用いて単語候補の一覧を作成しこれを出力装置に出力することを特徴とする。

【0013】本願の第8の発明は、第3及び第6の発明の前記修正用ナビゲーションとして入力された前記音声文章修正終了を指示する音声入力の場合は、前記音声記憶部に記憶された音声情報を削除し、前記出力装置に表示された文章のフォーマットを以前に入力された文章のフォーマットと同一にすることを特徴とする。

【0014】本願の第9の発明は、コンピュータディク

ーション装置において、ディクテーションモードにある時入力装置から音声入力された編集用ナビゲーションまたはディクテーションのみを受け取り修正モード時には前記入力装置から音声入力された修正用ナビゲーションのみを受け取る認識手段と、前記認識手段が前記編集用ナビゲーションを受け取った場合は改行、カレット移動などの前記編集用ナビゲーション種別に応じた処理を行いその結果を文章記憶部に記憶するとともに出力装置に出力する文章編集手段と、前記認識手段が前記ディクテーションを受け取った場合はその文字列を前記文章記憶部に記憶しその音声情報を音声記憶部に記憶し前記文字列を出力装置に出力し前記修正用ナビゲーションを受け取った場合は前記修正用ナビゲーション種別に応じた処理を行いその結果を前記出力装置に出力する文章修正手段とを含んで構成されることを特徴とする。

【0015】本願の第10の発明は、第9の発明における前記ディクテーションに対しては入力された音声に対応した文字列が、前記ナビゲーションに対しては対応する動作およびその動作が前記編集用ナビゲーションであるか前記修正用ナビゲーションであるかの情報が記憶されている認識辞書記憶部を予め具備し、前記認識手段は、与えられた前記音声入力が前記認識辞書記憶部に前記ナビゲーションとして登録されていればその情報により前記編集用ナビゲーションまたは前記修正用ナビゲーションであると判断し前記ナビゲーションとして登録されていなければ前記ディクテーションと判断することを特徴とする。

【0016】本願の第11の発明は、第9の発明における前記文章修正手段は、受け取った前記ディクテーションの文章を前記出力装置に出力する時、以前に入力された文章とフォーマットを変えて表示することにより、新しく入力された文章が区別できるようにさらにこの新しく入力された文章に対しての修正が可能な前記修正モードへの切り換えを行うことを特徴とする。

【0017】本願の第12の発明は、第9の発明における前記文章修正手段は、前記修正モードで入力された前記修正用ナビゲーションが単語候補の表示を指示する音声入力の場合は、前記認識手段に前記音声記憶部に記憶された該当単語の前記音声情報を渡して候補一覧の要求を行いその結果を前記出力装置に出力することを特徴とする。

【0018】本願の第13の発明は、第9及び11の発明における前記文章修正手段は、前記修正モードで入力された前記修正用ナビゲーションが文章修正終了を指示する音声入力の場合は、前記音声記憶部に記憶された音声情報を削除し、前記出力装置に表示されたフォーマットを以前に入力された文章のフォーマットと同一にすることを特徴とする。

【0019】本願の第14の発明は、データ処理装置にあって第9から第13の発明における少なくとも1つの

発明における前記認識手段及び前記文章編集手段及び前記文章修正手段と同等の動作を実行するコンピュータディクテーションプログラムを格納した記録媒体であることを特徴とする。

【0020】【作用】本発明は、ディクテーション機能およびナビゲーション機能を有する音声認識システムに関する。一発声分のディクテーション入力ごとに、自動的にディクテーションモードから修正モードへの切り換えを行い、逐次文章の修正を行うことにより、記憶領域の使用量を低減させ、音声認識の精度および操作性を向上させる方法を提供するものである。

【0021】ここで、ディクテーションとは、発声された音声进行を認識し、音声により文章を入力することをいう。また、ナビゲーションとは、あらかじめ登録された単語を音声認識し、音声により命令を実行することをいう。

【0022】以下の説明において、ディクテーションを受け付けている状態にある時をディクテーションモード、受け付けていない状態にある時を修正モードと定義し、ディクテーションモードで使用されるナビゲーションを編集用ナビゲーション、修正モードで使用されるナビゲーションを修正用ナビゲーションと定義する。修正用ナビゲーションは、新しくディクテーション入力された文章に対して、単語の候補の変更・決定、入力する文章の確定などの操作を行うことができる。編集用ナビゲーションは、以前に入力された文章に対して、文字入力位置（以後カレットと表現する）の移動、改行の挿入など、文字列の置換を行わない範囲での修正のみが可能である。また、ディクテーション結果とは、ディクテーション入力された音声情報と、音声に対応した文字列を示す。ここで音声情報とは、入力された音声の波形データおよび波形の解析結果のデータを表す。

【0023】図1において、認識コンポーネント21は、入力装置1から音声入力を受け取る。認識コンポーネント21は、与えられた入力を認識辞書記憶部31の情報を検索することにより、ディクテーション、編集用ナビゲーション、修正用ナビゲーションに分類する。認識辞書記憶部31には、ディクテーションおよびナビゲーションに関する情報が記憶されている。すなわち、ディクテーションに対しては入力された音声に対応した文字列が、ナビゲーションに対しては対応する動作および、その動作が編集用ナビゲーションであるか修正用ナビゲーションであるかの情報が記憶されている。認識コンポーネント21は、与えられた音声入力が認識辞書記憶部31にナビゲーションとして登録されていれば、その情報により編集用ナビゲーションまたは修正用ナビゲーションであると判断する。また、ナビゲーションとして登録されていなければ、ディクテーションと判断する。

【0024】初期状態では、データ処理装置2はディク

10

20

30

40

50

ーションモードにある。この時、認識コンポーネント 21 は、入力装置 1 から音声入力された編集用ナビゲーションまたはディクテーションのみを処理し、それ以外の入力を無視する。認識コンポーネント 21 が編集用ナビゲーションを受け取った場合は、文章編集コンポーネント 22 にこれを渡す。文章編集コンポーネント 22 は、改行、カレット移動などの編集用ナビゲーション種別に応じた処理を行い、結果を文章記憶部 32 に記憶し、出力装置 4 に出力する。

【0025】認識コンポーネント 21 がディクテーションを受け取った場合は、ディクテーション結果を文章修正コンポーネント 23 にこれを渡す。文章修正コンポーネント 23 は、文字列を文章記憶部 32 に記憶し、音声情報を音声記憶部 33 に記憶し、文章を出力装置 4 に出力する。この時、以前に入力された文章とフォントを変えるまたはフォントの色を変えて表示することにより、新しく入力された文章が区別できるようにする。さらに、この新しく入力された文章に対しての修正が可能な修正モードへの切り換えを、自動的に行う。

【0026】修正モード時には、認識コンポーネント 21 は、入力装置 1 から音声入力された修正用ナビゲーションのみを受け取り、文章修正コンポーネント 23 に渡す。文章修正コンポーネント 23 は、修正用ナビゲーション種別に応じた処理を行い、結果を出力装置 4 に出力する。単語候補の表示の場合は、認識コンポーネント 21 に該当単語の音声情報を渡して候補一覧の要求を行い、これを出力する。文章修正終了のナビゲーションを受け取ると修正モードは終了し、再びディクテーションモードに切り換わる。この時、修正を終えた文章は文章記憶部 32 に記憶され、音声情報は音声記憶部 33 から削除される。また、フォントまたは色を、通常の状態に戻して出力装置 4 に出力する。

【0027】修正用ナビゲーションを実行するためには、入力された音声情報が必要である。常に文章の修正を可能とするためには、入力された全ての音声情報を記憶しておく必要がある。そのためには、文字列のみを記憶する場合に比べ多くの記憶領域が必要となる。本発明では、一発声分のディクテーション入力ごとに自動的に修正モードに切り換え、入力された文章の修正を完了する。そのため、一発声ごとに音声情報を削除し、文字列のみを記憶することが可能であり、記憶領域の使用量を低減することができる。

【0028】また、ディクテーションモードと修正モードを切り換えることにより、編集用ナビゲーションおよび修正用ナビゲーションが使用できる状況を限定することができる。これにより、修正モード中に音声入力されたディクテーションおよび編集用ナビゲーションを無視することができ、修正モード中に入力された修正用ナビゲーションが、ディクテーションまたは編集用ナビゲーションと誤認識されることを防ぐ。同様に、ディクテ

ーションモード中に入力されたディクテーションまたは編集用ナビゲーションが、修正用ナビゲーションと誤認識されることを防ぐ。このため、音声認識の精度を向上させることができ、誤認識による誤動作が起こる可能性を低減できるため、操作性を向上させることが出来る。

【0029】

【発明の実施の形態】次に、本発明の実施の形態について図面を参照して詳細に説明する。

【0030】図 1 は、本発明の一実施の形態を示すブロック図である。

【0031】図 1 を参照すると、本実施の形態は、マイク等の入力装置 1 と、プログラム制御により動作するデータ処理装置 2 と、情報を記憶する記憶装置 3 と、ディスプレイ装置や印刷装置等の出力装置 4 とを含む。

【0032】データ処理装置 2 は、認識コンポーネント 21 と、文章編集コンポーネント 22 と、文章修正コンポーネント 23 とを備える。

【0033】認識コンポーネント 21 は、入力装置 1 から与えられた音声入力を、認識辞書記憶部 31 の検索により、ディクテーション、編集用ナビゲーション、修正用ナビゲーションに分類する。また、文章編集コンポーネント 22 より音声入力終了通知があった場合には処理を終了し、文章修正コンポーネント 23 より要求があった場合には、単語候補の一覧を作成する。音声入力に対して、ディクテーションモード中または修正モード中に応じて、次のような動作をする。

【0034】① ディクテーションモード中の場合

①-A 入力編集用ナビゲーションの場合

文章編集コンポーネント 22 に、編集用ナビゲーションを渡す。

【0035】①-B 入力ディクテーションの場合

文章修正コンポーネント 23 に、ディクテーション結果を渡す。

【0036】①-C 入力修正用ナビゲーションの場合

入力を無視し、何も行わない。

【0037】② 修正モード中の場合

②-A 入力編集用ナビゲーションの場合

入力を無視し、何も行わない。

【0038】②-B 入力ディクテーションの場合

入力を無視し、何も行わない。

【0039】②-C 入力修正用ナビゲーションの場合

文章修正コンポーネント 23 に、修正用ナビゲーションを渡す。

【0040】文章編集コンポーネント 22 は、認識コンポーネント 21 から編集用ナビゲーションを受け取る。

受け取ったナビゲーションが音声入力終了ナビゲーションであった場合は、その旨を認識コンポーネントに通知して処理を終了する。それ以外の場合はナビゲーション

を実行し、結果を文章記憶部32に記憶し、出力装置4に出力する。編集用ナビゲーションには、音声入力終了、改行、空白挿入、文章削除、カレット移動等がある。

【0041】文章修正コンポーネント23は、認識コンポーネント21から受け取った内容により、次のような動作をする。

【0042】① ディクテーション結果の場合
文章記憶部32に文字列を記憶し、音声記憶部33に音声情報を記憶し、出力装置4に出力する。また、修正モードへ切り換えを行う。

【0043】② 修正用ナビゲーションの場合
②-A 修正モード終了のナビゲーションの場合
ディクテーションモードへ切り換えを行う。また、修正を完了した文章を文章記憶部32に記憶し、音声記憶部33の内容を削除する。

【0044】②-B 候補一覧の表示ナビゲーションの場合

該当する単語の音声情報を認識コンポーネント21に渡し、候補の一覧を要求する。認識コンポーネント21が返した候補一覧を、出力装置4に出力する。

【0045】②-C 上記以外のナビゲーションの場合
修正用ナビゲーションを実行し、結果を出力装置4に出力する。修正用ナビゲーションには、現在選択されている文字列の変更、選択されている候補の変更、候補の決定等がある。次に、記憶装置3は、認識辞書記憶部31と、文章記憶部32と、音声記憶部33とを備える。

【0046】認識辞書記憶部31は、ディクテーションに対しては入力された音声に対応した文字列の情報を、ナビゲーションに対しては入力された音声に対応する動作および、その動作が編集用ナビゲーションであるか修正用ナビゲーションであるかの情報を記憶する。

【0047】文章記憶部32は、ディクテーションモードにおいて、文章編集コンポーネント22により編集された文章および文章修正コンポーネント23により入力される文章を記憶する。また、修正モードにおいて、文章修正コンポーネント23により修正を完了した文章を記憶する。

【0048】音声記憶部33は、文章修正コンポーネント23から渡される、ディクテーションの音声情報を記憶する。

【0049】次に、本発明の実施の形態の動作について図面を参照して説明する。

【0050】図2から図4はそれぞれ、本発明の認識コンポーネント21、文章編集コンポーネント22、文章修正コンポーネント23の処理フローである。

【0051】まず、図2を参照して、認識コンポーネント21の動作について説明する。初期状態において、データ処理装置2はディクテーションモードにあり(図2のステップA1)、認識コンポーネント21は音声入力

待ち状態になっている(ステップA2)。

【0052】入力装置1から与えられた音声入力は、認識コンポーネント21に供給される(ステップA3)。この時、認識コンポーネント21は、データ処理装置2がディクテーションモードにあるか否かを調べる(ステップA4)。

【0053】ディクテーションモードであった場合は、認識コンポーネント21は、入力された音声認識辞書記憶部31にナビゲーションとして登録されているか否かを調べる。登録されていない場合はディクテーション結果と判断し、これを文章修正コンポーネントに渡す(ステップA5)。登録されている場合は、さらに、編集用ナビゲーションとして登録されているか否かを調べる(ステップA6)。編集用ナビゲーションでなかった場合は、処理を行わず音声入力待ちになる(ステップA2)。編集用ナビゲーションであった場合は、これを文章編集コンポーネント22に渡し処理する(ステップA7)。文章編集コンポーネント22の処理終了後、認識コンポーネント21に入力終了通知があったか否かを調べる(ステップA8)。通知があった場合は認識コンポーネント21の動作は終了し、通知がなかった場合は音声入力待ちになる(ステップA2)。

【0054】ディクテーションモードでなかった場合は、認識コンポーネント21は、入力された音声認識辞書記憶部31に修正用ナビゲーションとして登録されているか否かを調べる(ステップA9)。登録されていない場合は、処理を行わず音声入力待ちになる(ステップA2)。登録されている場合はこれを文章修正コンポーネント23に渡し処理する(ステップA10)。文章修正コンポーネント23の処理後、認識コンポーネント21は候補要求があったか否かを調べる(ステップA11)。要求があった場合は、認識コンポーネント21は候補一覧を作成し(ステップA12)、これを文章修正コンポーネント23に渡し処理する(ステップA10)。要求がなかった場合は、音声入力待ちになる(ステップA2)。

【0055】次に、図2および図3を参照して、文章編集コンポーネント22の動作について説明する。

【0056】認識コンポーネント21より編集用ナビゲーションが渡されたとき(ステップA7)、文章編集コンポーネント22は、音声入力終了ナビゲーションであるか否かを調べる(図3のステップB1)。音声入力終了ナビゲーションであった場合は、認識コンポーネント21に終了通知を送り(ステップB2)、処理を終了する。音声入力終了ナビゲーションでなかった場合は、編集用ナビゲーションを実行する(ステップB3)。編集結果を文章記憶部32に記憶し、出力装置4に表示し、処理を終了する(ステップB4およびB5)。

【0057】続いて、図2および図4を参照して、文章修正コンポーネント23の動作について説明する。

【0058】認識コンポーネント21より入力があったとき(ステップA10)、文章修正用コンポーネント23は、ディクテーション結果か否かを調べる(ステップC1)。ディクテーション結果であった場合は、文字列を文章記憶部32に記憶し(ステップC2)、渡された音声情報を音声記憶部33に記憶し(ステップC3)、これまでに入力された文章と異なるフォントまたは色で出力装置4に表示する(ステップC4)。さらに、修正モードに切り換えを行い、処理を終了する(ステップC5)。

【0059】入力がディクテーション結果でなかった場合は、文章修正用コンポーネント23は、修正モード終了ナビゲーションか否かを調べる(ステップC6)。修正モード終了ナビゲーションであった場合は、修正を終えた文章を文章記憶部32に記憶し、音声記憶部33に記憶された音声情報を削除する(ステップC7およびC8)。また、文章のフォントまたは色を通常のものにして出力装置4に出力する(ステップC9)。さらに、ディクテーションモードに切り換えを行い、処理を終了する(ステップC10)。

【0060】入力が修正モード終了ナビゲーションでなかった場合は、文章修正用コンポーネント23は、候補一覧の表示ナビゲーションか否かを調べる(ステップC11)。候補一覧の表示ナビゲーションであった場合は、認識コンポーネント21に候補要求を出し(ステップC12)、得られた候補一覧を出力装置4に出力する(ステップA11、A12およびC13)。候補一覧の表示ナビゲーションでなかった場合は、ナビゲーションを実行する(ステップC14)。

【0061】次に、具体例を用いて本実施の形態の動作を説明する。

【0062】認識辞書記憶部31には、編集用ナビゲーションとして、以下のものが登録されているとする。

①「ここにスペース」現在のカレット位置に空白を挿入する。

②「編集終了」音声入力を終了する。音声入力終了ナビゲーション。

【0063】また、修正用ナビゲーションとして、以下のものが登録されているとする。

①「候補の表示」現在選択されている単語の修正候補を表示する、候補一覧の表示ナビゲーション。

②「候補の選択」候補一覧が表示されている時に、候補が選択状態になっていなければ先頭の候補を選択状態にし、選択状態になっているときはその次の候補を選択状態にする。結果を出力装置4に出力する。

③「候補の決定」候補一覧が表示されている時に、現在選択されている候補に確定する。候補一覧表示を終了させ、現在の選択位置の文字列を選択された候補に置換し、出力装置4に出力する。

④「修正終了」文章の修正を終了し、修正モードからデ

ィクテーションモードに切り換える。また、修正を完了した文章を黒色で、修正モードにある文章を赤色で表示するものとする。

【0064】初期状態では、データ処理装置2はディクテーションモードにあり、入力待ち状態となっている(ステップA1およびA2)。ここで、入力装置1より「ここにスペース」の音声が入力されたとする(ステップA3)。この時、認識コンポーネント21は、現在ディクテーションモードか否かを調べる(ステップA4)。ここではディクテーションモードと設定されているため、認識コンポーネント21は次に、入力された音声認識辞書記憶部31にナビゲーションとして登録されているか否かを調べる。登録されていない場合はディクテーションと判断する(ステップA5)。本実施例では、「ここにスペース」は編集用ナビゲーションとして登録されており、現在のカレット位置に空白を挿入すると定義されている。編集用ナビゲーションの入力であったため、認識コンポーネント21は文章編集コンポーネント22に、「ここにスペース」の編集用ナビゲーションを渡す(ステップA6およびA7)。文章編集コンポーネント22は、編集終了ナビゲーションではないと判断し、編集動作を実行する(ステップB1およびステップB3)。現在のカレット位置に空白を挿入し、文章記憶部32に記憶する(ステップB4)。さらに、その結果を出力装置4に出力する(ステップB5)。認識コンポーネント21は、文章編集コンポーネント22から入力終了通知がないことから、入力待ち状態になる(ステップA8およびA2)。

【0065】次に、「候補の表示」という音声入力があったとする(ステップA3)。現在ディクテーションモードであるため(ステップA4)、認識コンポーネント21は認識辞書記憶部31を調べ、ナビゲーションとして登録されているためディクテーションではないことがわかる(ステップA5)。しかし、編集用ナビゲーションとして登録されていないことから、この入力を無視して入力待ち状態になる(ステップA6およびステップA2)。

【0066】続いて、「今日はよい天気です」という音声入力をしようとしたが、「京はよい天気です」と誤認識されたとする(ステップA3)。現在ディクテーションモードであるため(ステップA4)、認識コンポーネント21は認識辞書記憶部31を調べ、ナビゲーションとして登録されていないことから「京はよい天気です」というディクテーション入力があったと判断し(ステップA5)、これを文章修正コンポーネント23に渡す(ステップA10)。文章修正コンポーネント23は、ディクテーション結果の入力であることから文章記憶部32に文字列を記憶し(ステップC1およびステップC2)、音声情報を音声記憶部33に記憶する(ステップC3)。また、「京はよい天気です」という文字列を、

10

20

30

40

50

修正可能なことを示すために赤色で出力装置4に表示する(ステップC4)。さらに、この文字列を修正するための修正モードへの切り換えを行い(ステップC5)、処理を終了する。認識コンポーネント21は、文章修正コンポーネント23から候補要求がないことから、入力待ち状態になる(ステップA11およびA2)。

【0067】ここで、「ここにスペース」の入力を行ったとする(ステップA3)。現在修正モードであるため、認識コンポーネント21は認識辞書記憶部31に修正用ナビゲーションとして登録されているか否かを調べる(ステップA4およびステップA9)。その結果、修正用ナビゲーションとして登録されていないことから、これを無視して入力待ちになる(ステップA2)。同様に、「今日はよい天気です」の入力を再度行った場合(ステップA3)、現在修正モードであり、修正用ナビゲーションとして登録されていないことから(ステップA4およびA9)、これを無視して入力待ちになる(ステップA2)。

【0068】次に、「京」の文字を「今日」に修正するために、修正候補の表示を行う。今、「京」の文字が選択状態にあるとする。「候補の表示」の音声入力があると(ステップA3)、現在修正モードであるため、認識コンポーネント21は認識辞書記憶部31に修正用ナビゲーションとして登録されているか否かを調べる(ステップA4およびステップA9)。その結果、修正用ナビゲーションとして登録されていることから、文章修正コンポーネント23に渡す(ステップA10)。文章修正コンポーネント23は、「候補の表示」がディクテーション結果ではなく、修正モード終了ナビゲーションでもなく、候補一覧の表示ナビゲーションであることから(ステップC1、ステップC6およびステップC11)、認識コンポーネント21に現在選択されている「京」の音声情報を渡し、候補の一覧を要求する(ステップC12)。認識コンポーネント21は、文章修正コンポーネント23からの候補要求を受けて候補一覧を作成し(ステップA11およびステップA12)、これを文章修正コンポーネント23に渡す(ステップA10)。文章修正コンポーネント23は、受け取った候補一覧を出力装置4に出力し(ステップC13)、処理を終了する。認識コンポーネント21は、入力待ち状態になる(ステップA11およびステップA2)。この時、「候補の選択」の音声入力を行うと(ステップA3)、現在修正モードであり、修正用ナビゲーションであるため(ステップA4およびステップA9)、文章修正コンポーネント23に渡される(ステップA10)。文章修正コンポーネント23は、「候補の選択」がディクテーション結果ではなく、修正モード終了ナビゲーションでもなく、候補一覧の表示ナビゲーションでもないことから(ステップC1、ステップC6およびステップC11)、ナビゲーションを実行する(ステップC14)。

「今日」の文字列が候補の先頭にあったとすると、「今日」を選択状態にする。文章修正コンポーネント23の処理は終了し、認識コンポーネント21は、入力待ち状態になる(ステップA11およびステップA2)。ここで、「候補の決定」の音声入力があると、同様にナビゲーションが実行され、候補一覧表示を終了して出力装置4の「京」の文字を「今日」に置換する。

【0069】さらに「修正終了」の音声入力があった場合(ステップA3)、現在修正モードであり、修正用ナビゲーションであるため(ステップA4およびステップA9)、文章修正コンポーネント23に渡される(ステップA10)。文章修正コンポーネント23は、「修正終了」がディクテーション結果ではなく、修正モード終了ナビゲーションであるため(ステップC1およびステップC6)、文章記憶部に修正された文字列「今日はよい天気です」を記憶する(ステップC7)。また、音声記憶部に記憶された音声情報を削除し、出力装置4に赤色で表示されている「今日はよい天気です」を、黒色で再表示する(ステップC8およびステップC9)。その後、ディクテーションモードに切り換え(ステップC10)、処理を終了する。認識コンポーネント21は、入力待ち状態になる(ステップA11およびステップA2)。

【0070】再びディクテーションモードになったことにより、ディクテーションおよび「ここにスペース」などの編集用ナビゲーションが有効になり、「候補の表示」などの修正用ナビゲーションは無効になる。ここで、「編集終了」の音声入力があると(ステップA3)、現在ディクテーションモードであるため、認識コンポーネント21は認識辞書記憶部31を調べ、ナビゲーションとして登録されていることがわかる(ステップA4およびステップA5)。「編集終了」は編集用ナビゲーションであるので、これを文章編集コンポーネント22に渡す(ステップA6およびステップA7)。文章編集コンポーネント22は、「編集終了」が音声入力終了ナビゲーションか否かを調べる(ステップB1)。本実施例では、「編集終了」は編集用ナビゲーションであるので、文章編集コンポーネント22は認識コンポーネント21に入力終了通知を行い(ステップB2)、処理を終了する。認識コンポーネント21は入力終了通知があったため、処理を終了する(ステップA8)。

【0071】次に、本発明の他の実施の形態について図面を参照して説明する。

【0072】まず図5を参照して、第2の実施の形態を説明する。

【0073】本実施の形態では、文章修正コンポーネント73は、図1における文章修正コンポーネント23が音声情報を音声記憶部33に記憶する代わりに、単語候補の一覧を単語候補記憶部83に記憶する。データ処理装置7の認識コンポーネント71および文章編集コンポ

ーネット72は、認識コンポーネント21および文章編集コンポーネント22と同一の動作をする。また、記憶装置8の認識辞書記憶部81および文章記憶部82は、認識辞書記憶部21および文章記憶部22と同一の動作をする。文章修正コンポーネント23および73は、以下の3つの場合について異なる動作を行う。

【0074】第1の場合は、ディクテーションモードにおいてディクテーション結果を受け取ったときである。文章修正コンポーネント23は音声記憶部33に音声情報を記憶する(図4のステップC3)。認識コンポーネント73は、入力された全ての単語について単語候補の一覧を認識コンポーネント21に要求し、これを単語候補記憶部83に記憶し、音声情報を記憶しない。

【0075】第2の場合は、修正モードにおいて候補一覧の表示ナビゲーションを受け取った場合である。文章修正コンポーネント23は認識コンポーネント21に音声情報を渡して単語候補一覧を要求する(ステップC12)。文章修正コンポーネント73は、前述のディクテーションを受け取った場合に作成した候補一覧を、単語候補記憶部83に要求する。

【0076】第3の場合は、修正モードにおいて修正モード終了ナビゲーションを受け取った場合である。文章修正コンポーネント23は、音声記憶部33の情報を削除する(ステップC8)。文章修正コンポーネント73は、単語候補記憶部83の情報を削除する。

【0077】次に、図6を参照すると、第3の実施の形態は、ディクテーションプログラムを記録した記録媒体5を備える。この記録媒体5は磁気ディスク、半導体メモリその他の記録媒体であってよい。

【0078】ディクテーションプログラムは記録媒体5からデータ処理装置6に読み込まれ、データ処理装置6の動作を制御する。データ処理装置6はディクテーションプログラムの制御により以下の処理、すなわち図1に示された実施例におけるデータ処理装置2と同一の処理を実行する。

【0079】データ処理装置6は、入力装置1より与えられた入力を認識辞書記憶部31の情報を検索することにより、ディクテーション、編集用ナビゲーション、修正用ナビゲーションに分類する。認識辞書記憶部31に登録されていれば、その情報により編集用ナビゲーションまたは修正用ナビゲーションであると判断し、登録されていなければ、ディクテーションと判断する。

【0080】初期状態では、データ処理装置6はディクテーションモードにある。この時、入力装置1から音声入力された編集用ナビゲーションまたはディクテーション結果のみを処理し、それ以外の入力を無視する。編集用ナビゲーションが入力された場合は、改行、文章の削除、カレット移動などの編集用ナビゲーション種別に応じた処理を行い、結果を文章記憶記憶部32に記憶し、出力装置4に出力する。

【0081】入力装置1からディクテーション入力された場合は、データ処理装置6は、文字列を文章記憶部32に記憶し、音声情報を音声記憶部33に記憶し、文章を出力装置4に出力する。この時、以前に入力された文章とフォントを変えるまたはフォントの色を変えて表示することにより、新しく入力された文章が区別できるようにする。さらに、この新しく入力された文章に対しての修正が可能な、修正モードへの切り換えを行う。

【0082】修正モード時には、データ処理装置6は、入力装置1から音声入力された修正用ナビゲーションのみを受け取り、単語の選択、単語候補の表示などの修正用ナビゲーション種別に応じた処理を行い、結果を出力装置4に出力する。文章修正終了のナビゲーションを受け取ると修正モードは終了し、再びディクテーションモードに切り換わる。この時、修正を終えた文章は文章記憶部32に記憶され、音声情報は音声記憶部33から削除される。また、フォントまたは色を、通常の状態に戻して出力装置4に出力する。

【0083】尚、詳述はしないが、第4の実施の形態として、第3の実施の形態におけるデータ処理装置6が、図5に示された実施例におけるデータ処理装置7と同一の処理を実行する場合が考えられることはいうまでもない。

【0084】

【発明の効果】第1の効果は、記憶領域の使用量の低減である。その理由は、ディクテーション入力があった場合に自動的に修正モードへの切り換えを行い、一発声ごとに文章を修正し、逐次音声情報を削除し、文字列のみを記憶するためである。一般に、音声情報は文字列よりも多くの情報を含む。入力された全ての音声情報を保持する場合と逐次削除する場合では、長い文章を入力するほど使用する記憶容量の差は大きくなる。

【0085】第2の効果は、音声認識率の向上と、それに伴う操作性の向上である。その理由は、ディクテーションモードではディクテーションおよび編集用ナビゲーションのみが可能であり、修正モードでは修正用ナビゲーションのみが可能であるためである。これにより、与えられた音声入力が、ディクテーションモード中に修正用ナビゲーションと誤認識される、あるいは、修正モード中にディクテーションまたは編集用ナビゲーションと誤認識されることを防ぐことができる。本発明では、ディクテーション入力があった場合に自動的に修正モードへの切り換えを行い、修正用ナビゲーションのみが有効になる。このため、音声認識率が向上し、誤認識による誤動作と、その修正の手間を低減し、操作性が向上する。

【図面の簡単な説明】

【図1】本発明の一実施の形態を示すブロック図である。

【図2】本発明の認識コンポーネント21の処理フロー

図である。

【図3】本発明の文章編集コンポーネント22の処理フロー図である。

【図4】本発明の文章修正コンポーネント23の処理フロー図である。

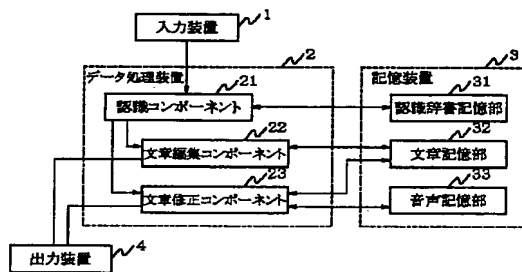
【図5】本発明の第2の実施の形態を示すブロック図である。

【図6】本発明の第3の実施の形態を示すブロック図である。

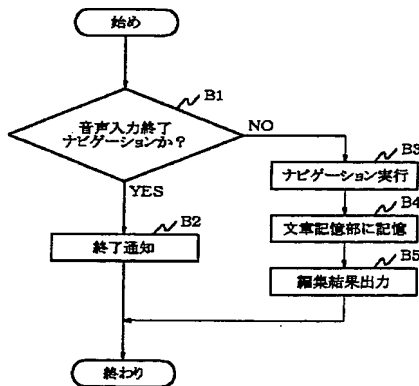
【符号の説明】

- 1 入力装置
- 2 データ処理装置
- 3 記憶装置
- 4 出力装置
- 5 記録媒体

【図1】

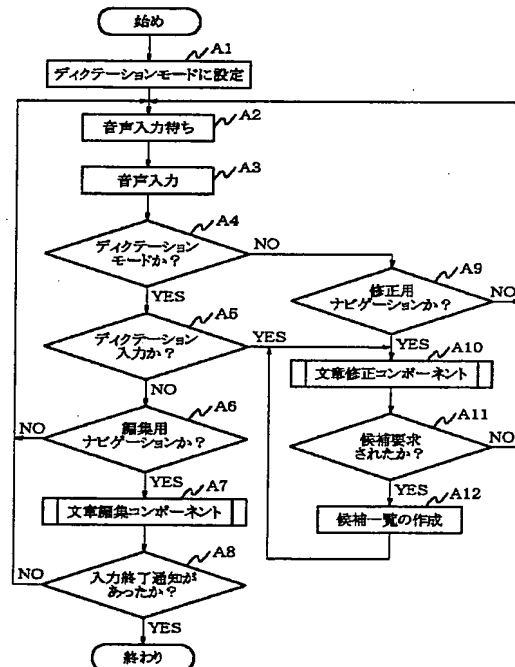


【図3】

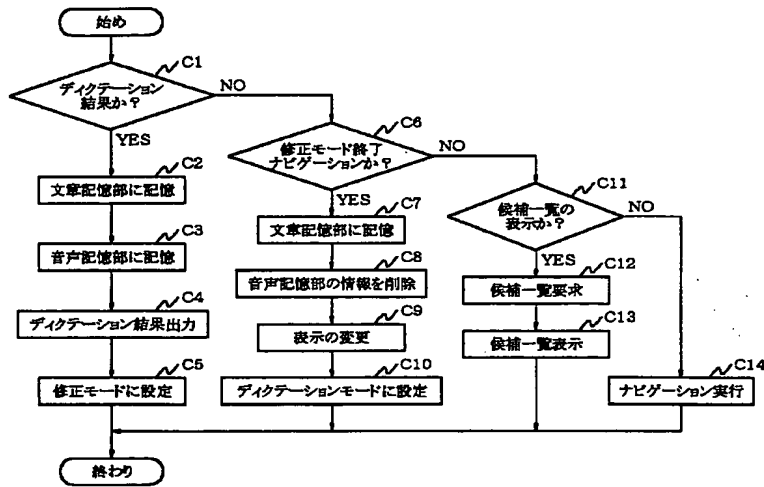


- 6 データ処理装置
- 7 データ処理装置
- 8 記憶装置
- 21 認識コンポーネント
- 22 文章編集コンポーネント
- 23 文章修正コンポーネント
- 31 認識辞書記憶部
- 32 文章記憶部
- 33 音声記憶部
- 10 71 認識コンポーネント
- 72 文章編集コンポーネント
- 73 文章修正コンポーネント
- 81 認識辞書記憶部
- 82 文章記憶部
- 83 単語候補記憶部

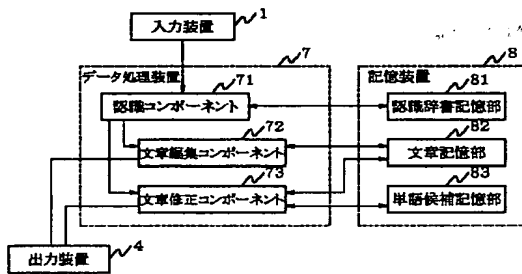
【図2】



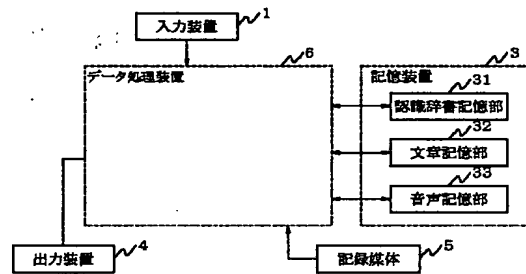
【図4】



【図5】



【図6】



THIS PAGE BLANK (USPTO)